

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日
Date of Application: 2 0 0 4 年 7 月 9 日

出 願 番 号
Application Number: 特 願 2 0 0 4 - 2 0 3 1 9 9

パリ条約による外国への出願
に用いる優先権の主張の基礎
となる出願の国コードと出願
番号

The country code and number
of your priority application,
to be used for filing abroad
under the Paris Convention, is

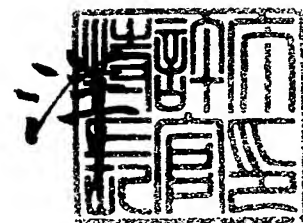
J P 2 0 0 4 - 2 0 3 1 9 9

出 願 人
Applicant(s): 日 本 電 信 電 話 株 式 有 限 公 司

2 0 0 5 年 8 月 3 日

特許庁長官
Commissioner,
Japan Patent Office

小 川



【書類名】	付 訂 願
【整理番号】	NTTH157579
【提出日】	平成16年 7月 9日
【あて先】	特許庁長官 殿
【国際特許分類】	H04B 10/00
【発明者】	
【住所又は居所】	東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
【氏名】	黒住 隆行
【発明者】	
【住所又は居所】	東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
【氏名】	永野 秀尚
【発明者】	
【住所又は居所】	東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
【氏名】	柏野 邦夫
【特許出願人】	
【識別番号】	000004226
【氏名又は名称】	日本電信電話株式会社
【代理人】	
【識別番号】	100064908
【弁理士】	
【氏名又は名称】	志賀 正武
【選任した代理人】	
【識別番号】	100108453
【弁理士】	
【氏名又は名称】	村山 靖彦
【手数料の表示】	
【予納台帳番号】	008707
【納付金額】	16,000円
【提出物件の目録】	
【物件名】	特許請求の範囲 1
【物件名】	明細書 1
【物件名】	図面 1
【物件名】	要約書 1
【包括委任状番号】	0401166

【請求項1】

蓄積音響信号から、該蓄積音響信号よりも短い乃至同じ長さの目的音響信号に類似した部分を探し出す信号検出システムにおいて、

前記蓄積音響信号の時系列データから、特徴ベクトルからなる蓄積特徴を計算する蓄積特徴計算部と、

前記目的音響信号の時系列データから、特徴ベクトルからなる目的特徴を計算する目的特徴計算部と、

前記蓄積特徴と、該蓄積特徴のサンプリング時間近傍における他の蓄積特徴とから、所定の統計量を計算し、蓄積特徴におけるベクトルの各要素ごとに正規化し、この正規化して得られた数値を要素とするベクトルからなる蓄積正規化特徴を導く蓄積特徴正規化部と

、
前記目的特徴と、該目的特徴のサンプリング時間近傍における他の目的特徴とから、所定の統計量を計算し、目的特徴におけるベクトルの各要素ごとに正規化し、この正規化して得られた数値を要素とするベクトルからなる目的正規化特徴を導く目的特徴正規化部と

、
前記蓄積正規化特徴から、量子化して得た値を要素とするベクトルからなる蓄積量子化特徴を導く蓄積特徴量子化部と、

前記目的正規化特徴から、量子化して得た値を要素とするベクトルからなる目的量子化特徴を導く目的特徴量子化部と、

前記蓄積量子化特徴中に照合区間を設定し、前記目的量子化特徴及び前記蓄積量子化特徴中の該照合区間のそれぞれとの類似度を計算する特徴照合部と

を有し、

前記特徴照合部が前記蓄積量子化特徴において照合区間を順次移動させて繰り返し、前記類似度から目的量子化特徴と類似する蓄積量子化特徴の領域を探索することを特徴とする音響信号検出システム。

【請求項2】

前記特徴ベクトルが所定の間隔毎にサンプリングされた周波数毎の強度情報を要素としたベクトルであり、前記統計量が照合区間における特徴ベクトルの平均値及び分散であることを特徴とする請求項1に記載の音響信号システム。

【請求項3】

前記目的特徴量子化部および前記蓄積特徴量子化部が、前記正規化特徴の各要素を、所定の閾値によりスカラー量子化して得た二値ベクトルを量子化特徴とする請求項1または請求項2に記載の音響信号検出システム。

【請求項4】

前記目的特徴量子化部および前記蓄積特徴量子化部が、前記正規化特徴のベクトルの要素のうち複数の要素をベクトル量子化することで得られた符号を要素に持つベクトルを量子化特徴とすることを特徴とする請求項1から請求項3のいずれかに記載の音響信号検出システム。

【請求項5】

目的音響信号に類似した蓄積信号の探索を要求するユーザ端末と、蓄積音響信号から、該蓄積音響信号よりも短い乃至同じ長さの前記目的音響信号に類似した部分を探し出す信号検出サーバとからなる信号検出システムにおいて、

前記ユーザ端末が、

前記目的音響信号の時系列データから、特徴ベクトルからなる目的特徴を計算する目的特徴計算部と、

前記目的特徴と、該目的特徴のサンプリング時間近傍における他の目的特徴とから、所定の統計量を計算し、目的特徴におけるベクトルの各要素ごとに正規化し、この正規化して得られた数値を要素とするベクトルからなる目的正規化特徴を導く目的特徴正規化部と

前記目的正規化特徴が、量子化して得た値を要素とするベクトルからなる蓄積特徴を導く蓄積特徴量子化部と、

を有し、

前記信号検出サーバが、

前記蓄積音響信号の時系列データから、特徴ベクトルからなる蓄積特徴を計算する蓄積特徴計算部と、

前記蓄積特徴と、該蓄積特徴のサンプリング時間近傍における他の蓄積特徴とから、所定の統計量を計算し、蓄積特徴におけるベクトルの各要素ごとに正規化し、この正規化して得られた数値を要素とするベクトルからなる蓄積正規化特徴を導く蓄積特徴正規化部と

前記蓄積正規化特徴から、量子化して得た値を要素とするベクトルからなる蓄積量子化特徴を導く蓄積特徴量子化部と、

前記蓄積量子化特徴中に照合区間を設定し、前記目的量子化特徴及び前記蓄積量子化特徴中の該照合区間のそれぞれとの類似度を計算する特徴照合部と

を有し、

前記特徴照合部が前記蓄積量子化特徴において照合区間を順次移動させて繰り返し、前記類似度から目的量子化特徴と類似する蓄積量子化特徴の領域を探索することを特徴とする音響信号検出システム。

【請求項6】

蓄積音響信号から、該蓄積音響信号よりも短い乃至同じ長さの、ユーザ端末から入力される目的音響信号に類似した部分を探し出す信号検出サーバにおいて、

前記目的音響信号の時系列データから、特徴ベクトルからなる目的特徴を計算する目的特徴計算部と、前記目的特徴と、該目的特徴のサンプリング時間近傍における他の目的特徴とから、所定の統計量を計算し、目的特徴におけるベクトルの各要素ごとに正規化し、この正規化して得られた数値を要素とするベクトルからなる目的正規化特徴を導く目的特徴正規化部と、前記目的正規化特徴から、量子化して得た値を要素とするベクトルからなる目的量子化特徴を導く目的特徴量子化部とを有するユーザ端末から前記目的量子化特徴を入力するユーザ信号入力部と、

前記蓄積音響信号の時系列データから、特徴ベクトルからなる蓄積特徴を計算する蓄積特徴計算部と、

前記蓄積特徴と、該蓄積特徴のサンプリング時間近傍における他の蓄積特徴とから、所定の統計量を計算し、蓄積特徴におけるベクトルの各要素ごとに正規化し、この正規化して得られた数値を要素とするベクトルからなる蓄積正規化特徴を導く蓄積特徴正規化部と

前記蓄積正規化特徴から、量子化して得た値を要素とするベクトルからなる蓄積量子化特徴を導く蓄積特徴量子化部と、

前記蓄積量子化特徴中に照合区間を設定し、前記目的量子化特徴及び前記蓄積量子化特徴中の該照合区間のそれぞれとの類似度を計算する特徴照合部と

を有し、

前記特徴照合部が前記蓄積量子化特徴において照合区間を順次移動させて繰り返し、前記類似度から目的量子化特徴と類似する蓄積量子化特徴の領域を探索することを特徴とする音響信号検出サーバ。

【請求項7】

蓄積音響信号から、該蓄積音響信号よりも短い乃至同じ長さの目的音響信号に類似した部分を探し出す信号検出方法において、

前記蓄積音響信号の時系列データから、特徴ベクトルからなる蓄積特徴を計算する蓄積特徴計算過程と、

前記目的音響信号の時系列データから、特徴ベクトルからなる目的特徴を計算する目的特徴計算過程と、

前記蓄積特徴と、該蓄積特徴のサンプリング時間近傍における他の蓄積特徴とから、所

定の統計量を計算し、蓄積特徴における、ベクトルの各要素ごとに正規化し、この正規化して得られた数値を要素とするベクトルからなる蓄積正規化特徴を導く蓄積特徴正規化過程と、

前記目的特徴と、該目的特徴のサンプリング時間近傍における他の目的特徴とから、所定の統計量を計算し、目的特徴におけるベクトルの各要素ごとに正規化し、この正規化して得られた数値を要素とするベクトルからなる目的正規化特徴を導く目的特徴正規化過程と、

前記蓄積正規化特徴から、量子化して得た値を要素とするベクトルからなる蓄積量子化特徴を導く蓄積特徴量子化過程と、

前記目的正規化特徴から、量子化して得た値を要素とするベクトルからなる目的量子化特徴を導く目的特徴量子化過程と、

前記蓄積量子化特徴中に照合区間を設定し、前記目的量子化特徴及び前記蓄積量子化特徴中の該照合区間のそれぞれとの類似度を計算する特徴照合過程と

を有し、

前記特徴照合過程において、前記蓄積量子化特徴において照合区間を順次移動させて繰り返し、前記類似度から目的量子化特徴と類似する蓄積量子化特徴の領域を探索することを特徴とする音響信号検出方法。

【請求項 8】

蓄積音響信号から、該蓄積音響信号よりも短い乃至同じ長さの目的音響信号に類似した部分を探し出す信号検出処理を実行するプログラムであり、

前記蓄積音響信号の時系列データから、特徴ベクトルからなる蓄積特徴を計算する蓄積特徴計算処理と、

前記目的音響信号の時系列データから、特徴ベクトルからなる目的特徴を計算する目的特徴計算処理と、

前記蓄積特徴と、該蓄積特徴のサンプリング時間近傍における他の蓄積特徴とから、所定の統計量を計算し、蓄積特徴におけるベクトルの各要素ごとに正規化し、この正規化して得られた数値を要素とするベクトルからなる蓄積正規化特徴を導く蓄積特徴正規化処理と、

前記目的特徴と、該目的特徴のサンプリング時間近傍における他の目的特徴とから、所定の統計量を計算し、目的特徴におけるベクトルの各要素ごとに正規化し、この正規化して得られた数値を要素とするベクトルからなる目的正規化特徴を導く目的特徴正規化処理と、

前記蓄積正規化特徴から、量子化して得た値を要素とするベクトルからなる蓄積量子化特徴を導く蓄積特徴量子化処理と、

前記目的正規化特徴から、量子化して得た値を要素とするベクトルからなる目的量子化特徴を導く目的特徴量子化処理と、

前記蓄積量子化特徴中に照合区間を設定し、前記目的量子化特徴及び前記蓄積量子化特徴中の該照合区間のそれぞれとの類似度を計算する特徴照合処理と

を有し、

前記特徴照合処理において、前記蓄積量子化特徴において照合区間を順次移動させて繰り返し、前記類似度から目的量子化特徴と類似する蓄積量子化特徴の領域を探索する処理をコンピュータに実行させるプログラム。

【請求項 9】

請求項 8 に記載の音響信号検出処理を行うプログラムを記録したコンピュータ読み取り可能な記録媒体。

【発明の名称】 音響信号検出システム及び音響信号検出サーバ

【技術分野】

【0001】

本発明は、蓄積された蓄積音響信号の中から、この蓄積音響信号よりも短い乃至同じ長さの目的音響信号と類似した信号の位置を探し出す信号検出に関するものであり、例えば実環境中の音響信号検出に用いられる音響信号検出システム及び音響信号検出サーバである。

すなわち、本発明は、実環境中に流れている音楽やCM（コマーシャル）を携帯端末で受信し、その受信された音響信号を用いて膨大な音楽CMデータベースの中から同一の音楽やCMを検索するものである。

【背景技術】

【0002】

近年、音や映像など、マルチメディア情報（本発明では例えば音楽）の流通が盛んに行われる様になり、このマルチメディア情報を得るための検索や探索が必要となってきた。

例えば、実環境中に流れている音楽やCMを携帯端末などにより受信し、その受信された音響信号（目的音響信号）を用いて、膨大な音楽CMデータベースの中から同一の音楽やCMを検索することが考えられる。

このため、具体的な音の目的音響信号を指定して、この目的音響信号と類似する蓄積音響信号が膨大なデータベースのいずれに存在するかの探索を行う必要がある。

例えば、高速信号検出法として、あらかじめ登録した音響信号と類似した音響信号の場所を探し出す音響信号検出方法が知られている（例えば、特許文献1参照）。

ここで行われる探索は、時系列探索であり、高速に、かつ高精度に行われる必要がある。

【0003】

ところが、ユーザが入手する音楽には実環境により収録されるものであるため、様々な特性ひずみ、例えば、スピーカなどの信号発生源の機器の特性、入力する携帯端末の特性による情報性ひずみや、実環境中の雑音、反響、吸収などによる加法性ひずみなどが、目的音響信号に含まれていることが考えられる。

しかしながら、特許文献1の検出法は、目的音響信号または蓄積音響信号のノイズによる特徴ひずみが少ないことが想定されて構成されており、雑音やひずみがある場合、探索精度が低下するという欠点がある。

【0004】

この欠点を解決するため、入力される音響信号に対する変動付加過程を設けることで特徴ひずみに対して頑健な信号検出を行う方法が提案されている（例えば、特許文献2参照）。

ところが上述した変動付加過程を設ける際、複数の雑音やひずみを考慮する必要がある場合、これらの複数の目的特徴を用意しなければならず、情報量が大幅に増大してしまうという欠点がある。

【0005】

また、入力される目的音響信号のパワーのピークを検出し、このピークにおける周波数を特徴とすることにより、ひずみに対して頑健な信号検出を行う方法が考案されている。

しかしながら、この方法には、入力される信号のピークを用いるため、この実際の信号のピーク周辺にある大きな雑音の影響を受けることにより、実際の信号のピーク検出に失敗し、探索精度が低下する欠点がある。

【0006】

このため、局所的な特徴の統計量を用いて、入力される目的音響信号を正規化することにより、ひずみに対して頑健な信号検出を行う方法が提案されている（例えば、特許文献3、4参照）。

りなわつ、この信号検出方法は、特徴ひずみによる目的音響信号の変動を吸収するため、周波数特徴を抽出した後、時間一周波数空間上の局所領域毎に目的音響信号を正規化して、特徴ひずみに頑健な空間へのデータ変換を行い、この空間において蓄積音響信号との比較を行う。

【特許文献1】特許第3065314号公報

【特許文献2】特許第3408800号公報

【特許文献3】特開2003-022084号公報

【特許文献4】WO 02/11123 A2

【発明の開示】

【発明が解決しようとする課題】

【0007】

しかしながら、上述した特徴ひずみに頑健な空間へのデータ変換を行う方法においては、非定常な雑音やひずみに対しては、正規化だけでひずみを吸収することができず、探索精度が低下するという欠点がある。

さらに、上述した各方法においては、探索精度とともに、データベースに記憶されている膨大な蓄積音響信号のなかから、目的音響信号に対応する蓄積音響信号を、細かな精度の数値データを比較して探索するため、多くの時間が必要となり、かつ探索精度を向上させるために、膨大な量の蓄積音響信号をデータベースに蓄積させる必要がある。

【0008】

本発明はこのような事情に鑑みてなされたものであり、検索対象の目的音響信号を、特徴ひずみに対して頑健なデータに変換して探索精度を向上させ、かつ探索時間を大幅に短縮させる音響信号検出システム及び音響信号検出サーバを提供することを目的とする。

【課題を解決するための手段】

【0009】

上記目的を達成するために、本発明の音響信号検出システムは、蓄積音響信号から、該蓄積音響信号よりも短い乃至同じ長さの目的音響信号に類似した部分を探し出す信号検出システムにおいて、前記蓄積音響信号の時系列データから、特徴ベクトルからなる蓄積特徴を計算する蓄積特徴計算部と、前記目的音響信号の時系列データから、特徴ベクトルからなる目的特徴を計算する目的特徴計算部と、前記蓄積特徴と、該蓄積特徴のサンプリング時間近傍における他の蓄積特徴とから、所定の統計量を計算し、蓄積特徴におけるベクトルの各要素ごとに正規化し、この正規化して得られた数値を要素とするベクトルからなる蓄積正規化特徴を導く蓄積特徴正規化部と、前記目的特徴と、該目的特徴のサンプリング時間近傍における他の目的特徴とから、所定の統計量を計算し、目的特徴におけるベクトルの各要素ごとに正規化し、この正規化して得られた数値を要素とするベクトルからなる目的正規化特徴を導く目的特徴正規化部と、前記蓄積正規化特徴から、量子化して得た値を要素とするベクトルからなる蓄積量子化特徴を導く蓄積特徴量子化部と、前記目的正規化特徴から、量子化して得た値を要素とするベクトルからなる目的量子化特徴を導く目的特徴量子化部と、前記蓄積量子化特徴中に照合区間を設定し、前記目的量子化特徴及び前記蓄積量子化特徴中の該照合区間のそれぞれとの類似度を計算する特徴照合部とを有し、前記特徴照合部が前記蓄積量子化特徴において照合区間を順次移動させて繰り返し、前記類似度から目的量子化特徴と類似する蓄積量子化特徴の領域を探索することを特徴とする。

【0010】

特に、「高速信号検出法、装置およびその記録媒体」（特許文献1）及び「信号検出方法、装置及びそのプログラム、記録媒体」（特許文献2）に比較して、新たに目的特徴正規化部、蓄積特徴正規化部、目的特徴量子化部及び蓄積特徴量子化部を設けて特徴ベクトルの各要素を、正規化した後、所定の閾値により量子化することで、ひずみに対して特許文献1及び2に比較して頑健な音響信号検出を可能とした。

【0011】

また、「信号検出方法及び装置、プログラムならびに記録媒体」（特許文献3）の目的

后々及び蓄積特徴量子化部を設け、特徴ベクトルの各要素を量子化することにより、特徴ひずみに対して上記手法に比較して頑健な音響信号検出を可能とした。

さらに、「System and methods for recognizing sound and music signal in high noise and distortion」(特許文献4)に比較して、新たに目的特徴正規化過程及び蓄積特徴正規化過程を設けて特徴を正規化することにより、特徴ひずみに対して上記手法に比較して頑健な音響信号検出を可能とした。

【発明の効果】

【0012】

本発明によれば、特徴ひずみを吸収するため、所定の閾値により音響信号をスカラー量子化する量子化部(過程)を設けることにより、比較するデータ量を大幅に低減することが可能となり、類似度の計算処理を高速化するとともに、蓄積する蓄積音響信号の1件あたりのデータ量を大幅に削減することができ、従来と同様の記憶部容量により多くの蓄積音響信号情報を蓄積することができ、この点からも類似した蓄積音響信号の検出精度を向上させることができる。

【0013】

また、本発明によれば、上記量子化部(過程)を設けることにより、各要素毎のデータの詳細な比較判定でなく、照合区間における特徴ベクトルを全体的なパターンとして検索するので、探索精度を向上させて、検出洩れを防ぐことが可能となり、公知の方法と比較して、さまざまな特徴ひずみに頑健な信号検出処理が行え、より汎用的な特徴ひずみに頑健な音響信号検出を行うことができる。

【発明を実施するための最良の形態】

【0014】

図1は、本発明の一実施例であり、音響信号を対象とする特徴ひずみに頑健な音響信号検出システムの構成を示すブロック図である。

この図1に示す音響信号検出システムは、音響信号を対象とする特徴ひずみに頑健な信号検出を実現するものであり、目的特徴計算部1と、蓄積特徴計算部2と、目的特徴正規化部3と、蓄積特徴正規化部4と、目的特徴量子化部5と、蓄積特徴量子化部6と、特徴照合部7と、蓄積量子化特徴データベース8とで構成され、蓄積時系列信号(蓄積音響信号)すなわち検索したい音響信号と、目的時系列信号(目的音響信号)すなわち検索される音響信号を入力し、目的時系列信号と類似した蓄積時系列信号中の箇所を出力する。

【0015】

目的特徴計算部1は、目的音響信号を所定の間隔で時系列にサンプリングして得られる離散値である目的時系列信号から、例えば、サンプリング単位に周波数毎のパワースペクトル値等を抽出し、この抽出した特徴量を多次元ベクトル化して特徴ベクトルを得て、この特徴ベクトルからなる目的特徴を導く(計算する)ものである。

蓄積特徴計算部2も、同様に、蓄積音響信号を所定の間隔で時系列にサンプリングして得られる離散値である蓄積時系列信号から、例えば、サンプリング単位に周波数毎のパワースペクトル値等を抽出し、この抽出した特徴量を多次元ベクトル化して特徴ベクトルを得て、この特徴ベクトルからなる蓄積特徴を導くものである。

【0016】

目的特徴正規化部3は、上記目的特徴から、隣接した部分も含む周辺の特徴から導いた統計量を用いて特徴ベクトルの各要素ごと独立に正規化し、正規化された値からなる多次元ベクトルを有する目的正規化特徴を導くものである。

蓄積特徴正規化部4は、上記蓄積特徴から、隣接した部分も含む周辺の蓄積特徴から導いた統計量を用いて特徴ベクトルの各要素ごと独立に正規化し、正規化された値からなる多次元ベクトルを有する蓄積正規化特徴を導くものである。

【0017】

目的特徴量子化部5は、上記目的正規化特徴を所定の閾値によりスカラー量子化し、この量子化により得られた値を要素とする多次元ベクトルからなる目的量子化特徴を導くも

っている。

蓄積特徴量子化部6は、上記蓄積正規化特徴を所定の閾値によりスカラー量子化し、この量子化により得られた値を要素とする多次元ベクトルからなる蓄積量子化特徴を導くものである。

【0018】

量子化を行う上記所定の閾値は、目的及び蓄積正規化特徴を、例えば2値化するとした場合、2値化して得られる要素の数値と、もとの要素の数値との平均2乗誤差の最小化される点を選択する等で求められる。

この目的量子化特徴及び蓄積量子化特徴は、各々時系列にサンプリングした順序に、各サンプリングにおける特徴ベクトルが順次配列している。配列数（配列の長さ）は蓄積量子化特徴より、目的量子化特徴の方が短いまたは等しい。

【0019】

特徴照合部7は、上記蓄積量子化特徴における時系列に配列している特徴ベクトルにおいて、この配列の所定の範囲を照合区間として設定し、この照合区間と同一の長さの被照合区間を目的量子化特徴に設定して、上記照合区間と被照合区間との類似度を計算し、予め設定された探索閾値と比較して、類似しているか否かの判定を行う。

また、特徴照合部7は、上記照合区間の照合処理が終了すると、新たな照合区間を設定するため、配列の隣接する同一時間幅の範囲にずらす。

蓄積量子化特徴データベース8は、複数の、例えば良く配信される曲に対して、蓄積特徴計算部2、蓄積特徴正規化部4及び蓄積特徴量子化部6により、予め蓄積量子化特徴を計算して、各曲名に対応して記憶されている。

【0020】

次に、図面を参照して本実施例の音響信号検出システムの動作の説明を行う。図2は図1の音響信号検出システムの動作例を示すフローチャートである。

蓄積特徴計算部2は、与えられた蓄積音響信号を読み込んで入力し（ステップS1）、入力した蓄積音響信号に対して特徴抽出を行う。

このとき、蓄積特徴計算部2は、音響信号のフーリエ変換の振幅成分を用い、例えば周波数8000Hzで標本化した音響信号の1秒の区間をフーリエ変換し、0～4000Hzを等間隔に32の周波数帯域の区間に分割し、0.1秒ごとに、各区間内での振幅成分の平均パワーからなる32次元の多次元ベクトルを特徴ベクトルとして抽出して蓄積特徴とする（ステップS2）。

【0021】

次に、蓄積特徴正規化部4は、蓄積特徴計算部2から蓄積特徴を読み込み、この蓄積特徴の特徴ベクトルの各要素毎に、所定の一定区画の平均値と標準偏差とを計算して求める。

例えば、蓄積特徴正規化部4は、各周波数帯域ごとにある時間区間の値から平均値と標準偏差とを求め、求められた平均値と標準偏差とを用いて正規化を行う。

このとき、蓄積特徴正規化部4による正規化後の周波数特徴 $y(i)$ のk番目の要素は、以下の（1）式となる。

【0022】

【数1】

$$y(i,k) = \frac{1}{\sigma(i,k)} (Q(i,k) - m(i,k)) \quad \dots(1)$$

【0023】

ただし、（1）式において、 $m(i,k)$ は平均値であり、以下に示す（2）式により求められ、 $\sigma(i,k)$ は標準偏差であり、以下に示す（3）式により求められる。

【0024】

$$m(i,k) = \frac{1}{2M} \sum_{i=-M}^{M-1} Q(i,k) \quad \dots(2)$$

【0025】

【数3】

$$\sigma(i,k)^2 = \frac{1}{2M} \sum_{i=-M}^{M-1} (Q(i,k) - m(i,k))^2 \quad \dots(3)$$

【0026】

(2) 式及び (3) 式において、M は局所時間内の周波数特徴の平均値及び標準偏差を求めるための時間窓の大きさの半分の数値である。

そして、(1) ~ (3) 式により得られた多次元ベクトルの時系列な配列 (k で示される順番) を蓄積正規化特徴とする (ステップ S3)。

【0027】

次に、蓄積特徴量子化部 6 は、蓄積特徴正規化部 4 から、蓄積正規化特徴を読み込んで入力し、蓄積正規化特徴の多次元ベクトルの各要素毎に、あらかじめ与えられた閾値を境に二値化処理をして量子化された数値のベクトルを得る。

例えば、蓄積特徴量子化部 6 は、蓄積正規化特徴の各要素を、得られた閾値 t を境に量子化する。

これにより、蓄積特徴量子化部 6 は、量子化後の周波数特徴 z(i) の k 番目の要素を、以下に示す (4) 式により各々計算する。

【0028】

【数4】

$$z(i,k) = \begin{cases} 1 & (y(i,k) > t \text{ のとき}) \\ 0 & (y(i,k) \leq t \text{ のとき}) \end{cases} \quad \dots(4)$$

【0029】

そして、蓄積特徴量子化部 6 は、各要素を 2 値化して、この多次元ベクトルを蓄積量子化特徴として出力する (ステップ S4)。

このとき、蓄積特徴量子化部 6 は、計算により得られた蓄積量子化特徴を、特徴照合部 7 へ直接送信するか、または一旦蓄積量子化特徴データベース 8 への登録のいずれかの処理を行う。

特徴照合部 7 がリアルタイムに蓄積量子化特徴と、目的量子化特徴との比較を行う場合、蓄積特徴量子化部 6 は特徴照合部 7 へ、入力されている蓄積音響信号の蓄積量子化特徴を出力し、蓄積量子化特徴データベース 8 へ蓄積音響信号のデータを登録する場合、蓄積特徴量子化部 6 は特徴照合部 7 へ蓄積量子化特徴を送信せずに、蓄積量子化特徴データベース 8 へ、曲名に対応させて蓄積量子化特徴の登録処理を行う。

【0030】

目的特徴計算部 1 は、与えられた目的音響信号を読み込んで入力し (ステップ S5)、入力した目的音響信号に対して特徴抽出を行う。

このとき、目的特徴計算部 1 は、蓄積特徴計算部 2 と同様に、音響信号のフーリエ変換の振幅成分を用い、例えば周波数 8000 Hz で標本化した音響信号の 1 秒の区間をフーリエ変換し、0 ~ 4000 Hz を等間隔に 32 の周波数帯域の区間に分割し、0.1 秒ご

次に、各区間内での振幅成分の平均パワーからなる多次元の多次元ベクトルを特徴ベクトルとして抽出して目的特徴とする（ステップS6）。

【0031】

次に、目的特徴正規化部3は、目的特徴計算部1から目的特徴を読み込み、この目的特徴の特徴ベクトルの各要素毎に、所定の一定区画の平均値と標準偏差とを計算して求める。

すなわち、目的特徴正規化部3は、蓄積特徴正規化部4と同様に、（1）～（3）式により得られた多次元ベクトルの時系列な配列（kで示される順番）を目的正規化特徴とする（ステップS7）。

【0032】

次に、目的特徴量子化部5は、目的特徴正規化部3から、目的正規化特徴を読み込んで入力し、目的正規化特徴の多次元ベクトルの各要素毎に、あらかじめ与えられた閾値を境に、蓄積特徴量子化部6と同様に、（4）式を用いて二値化処理をして、量子化された数値のベクトルを得る。

そして、目的特徴量子化部5は、各要素が2値化された、この多次元ベクトルを目的量子化特徴として出力する（ステップS8）。

【0033】

次に、特徴照合部7は、目的特徴量子化部5及び蓄積特徴量子化部6により得られた目的量子化特徴及び蓄積量子化特徴を各々読み込む。

また、特徴照合部7は、目的音響信号及び蓄積音響信号を同時に入力して、リアルタイムに類似を判定する場合以外、蓄積量子化特徴データベース8から順次比較を行う蓄積量子化特徴を読み出して、目的量子化特徴と比較する。

このとき、特徴照合部7は、蓄積量子化特徴において、目的特徴量子化部5で与えられた目的量子化特徴と同じ長さの特徴ベクトルの配列を照合区間として設定する。

【0034】

そして、特徴照合部7は、目的量子化特徴と、上記照合区間との類似度の計算を、類似度として、双方の特徴ベクトル間のハミング距離を計算して行う（ステップS9）。

例えば、特徴照合部7は、目的量子化特徴が15秒の長さであるとする、この目的量子化特徴の特徴ベクトルの配列から、配列の要素としての特徴ベクトルを、0.1秒間隔に合計150箇所抽出し、各サンプリングにおいて周波数帯域を32分割しているため、これらのベクトルからなる 150×32 の4800次元を照合に用いる目的ベクトルとする。

【0035】

また、特徴照合部7は、上記目的ベクトルと同様に、蓄積量子化特徴のデータの先頭から、15秒の長さを単位に順次照合区間として設定して、特徴ベクトルの配列から、特徴ベクトルを0.1秒間隔に合計150箇所抽出し、各サンプリングにおいて周波数帯域を32分割しているため、それらからなる 150×32 の4800次元を照合に用いる蓄積ベクトルとする。

【0036】

このとき、特徴照合部7が目的量子化特徴の特徴ベクトルの配列から、複数の箇所の要素を抽出して、目的ベクトルとするのであれば、目的特徴量子化部5において、予め配列の要素としての特徴ベクトルを目的正規化特徴から抽出して、すなわち、0.1秒間隔に合計150箇所抽出して、スカラー量子化を行い、目的ベクトルとして特徴照合部7へ出力するようにしてもよい。

【0037】

このとき、特徴照合部7は、照合区間を、蓄積量子化特徴の先頭から順次ずらしながら、目的量子化特徴との、すなわち上記目的ベクトルと蓄積ベクトルとのハミング距離を計算することにより比較処理を行う。

そして、特徴照合部7は、蓄積量子化特徴を最後まで照合処理した後、照合区間毎のハミング距離を参照して、ハミング距離が最も小さい照合区間の領域を探索結果として出力

る。

なお、特徴照合部 7 は、探索結果として、予めハミング距離の探索閾値を与えられていた場合、この探索閾値と選択された照合区間のハミング距離とを判定して（ステップ S 1 0）、この探索閾値を下回るもののみ出力することもできる（ステップ S 1 1）。

【0038】

また、特徴照合部 7 は、探索閾値を複数の照合区間のハミング距離が下回った場合、ハミング距離の上位（低い数値のものから）N 番目までの照合区間を出力するようにすることも可能である。

さらに、特徴照合部 7 は、探索閾値を下回る照合区間が無い場合など、該当箇所がないことを示す情報を通知し、新たな蓄積音響信号の蓄積量子化特徴を、蓄積量子化特徴データベース 8 から読み出し、上記探索閾値以下の照合区間を有する蓄積量子化特徴が探索されるまで、ステップ S 9 以降の探索処理を継続させて行うようにしてもよい。

【0039】

また、図 1 における目的特徴計算部 1、目的特徴正規化部 3 及び目的特徴量子化部 5 を各ユーザの端末（例えば、パーソナルコンピュータ）にインストールしておき、音楽配信を行うサービスプロバイダに蓄積特徴計算部 2、蓄積特徴正規化部 4、蓄積特徴量子化部 6、特徴照合部 7 及び蓄積量子化特徴データベース 8 を有する音響信号検出サーバを設けてもよい。

これにより、ユーザが携帯電話等で受信した音響信号を目的音響信号として、目的量子化特徴まで生成し、この目的量子化特徴を上記音響信号検出サーバへ、インターネットなどを介して送信し、この目的量子化特徴に類似した蓄積音響信号を探索してもらうように要求する構成とする。

【0040】

次に、上記音響信号検出システムを適用した場合の動作実験例及びその結果を示す。

本発明の音響信号検出システムによる効果を確認するため、本発明を適用した場合と適用しなかった場合との探索精度の比較を行った。

実験のため、蓄積音響信号として、ある C D（コンパクトディスク）の曲を再生した音響信号を、そのまま図 1 の音響信号検出システムの蓄積特徴計算部 2 に入力させた。

一方、目的音響信号として、スピーカにより、上記 C D の同一の曲の所定の部分（蓄積音響信号の一部）を鳴らし、その喫茶店内で携帯電話を P H S（登録商標）にかけ、その P H S（登録商標）の音声出力から受音して得た音響信号を、目的音響信号として、目的特徴計算部 1 へ入力させた。

【0041】

上述した同一の実験条件において、200 回繰り返して探索処理を行い、探索精度の測定を行った。

この探索精度は、探索閾値の設定を調節して、適合率（precision rate）と再現率（recall rate）とが等しくなる時の値とした。

ここで、適合率とは探索結果として出力されたもののうち正しいものの割合であり、再現率とは探索されるべきもののうち探索結果として出力されたものの割合である。

適合率または再現率は、探索閾値の設定によって変化するが、本実験においては探索閾値（ハミング距離に対する閾値）を以下に示す（5）式によって定めた。

【0042】

【数 5】

$$\theta = m + n\psi \quad \dots(5)$$

【0043】

上記（5）式において、m 及び ψ 各々は、与えられた蓄積信号に対して入力信号をサンプリングし、予備的に類似度の計算を行って収集した類似度の平均値と標準偏差であり、

11は聴感時に与えられる印象である。

ただし、(5)式において、探索閾値 θ が1を越える場合、 $\theta = 1$ とし、0を下回るときは $\theta = 0$ とした。

本実験においては、(5)式における n の値を200回繰り返して行う間で一定とし、その t の値を適合率と再現率とがほぼ等しくなるように調節した。

上述した実験の結果、上記精度は、正規化特徴の量子化を行わなかった場合に60.0%であり、量子化を行った場合(一実施例)で85.77%であった。

【0044】

また、探索処理においては、CPUにIntel(登録商標)Pentium(登録商標)III 1133MHz、OSにRedHat(登録商標)7.3、コンパイラにGNU gccを用いた。

また、実行ファイルはコンパイラ最適化オプション「-O3」によりコンパイルを行った。

上述した実験において、本発明の音響信号検出システムにより、探索精度が向上したことを確認することができる。

【0045】

本発明の音響信号検出システムは、実環境で収録した端末特性、符号化特性などによりひずみをうけた断片的な音響信号をつかって、一致する音楽やCMの曲名や放送時間を探索するだけでなく、音楽やCMに関連する情報を蓄積した付随情報データベースと組み合わせ情報探索するのにも利用できる。

例えば、TVやラジオから流れる音楽或いはCMをユーザが携帯電話で受信し、音響信号検索サービスなどに送信する。音響信号検索サービス提供者は、この音響信号に一致又は類似する音響をデータベースより検索し、音楽或いはこのCMに関する情報(例えば演奏者、作詞作曲者名、商品情報、製品特徴、サービス、撮影地、出演者、ホームページ等)をインターネットなどのネットワークを介して有料、または無料でユーザに提供するという構成が可能である。

【0046】

ここで、検索する音響の入力方法においては、携帯端末からサーバにアクセスし、端末のマイクで直接信号を受信し目的信号としてもかまわない。また、いったん録音した音響を適当に切り取って送信してもよい。

また、他の発明の実施例として、ユーザが家庭用テープレコーダで録音したイメージ音響から、そのイメージに近い音楽のフレーズを配信されているCDや映画、テレビプログラムから検索し、コンテンツを利用したり、また録音したイメージ音響に相当するフレーズを集めて編集したりすることにも利用できるであろう。

【0047】

なお、図1における音響信号検出システム及び音響信号検出サーバの機能を実現するためのプログラムをコンピュータ読み取り可能な記録媒体に記録して、この記録媒体に記録されたプログラムをコンピュータシステムに読み込ませ、実行することにより、音響信号検出処理及びデータベースに対する蓄積音響信号の蓄積処理を行ってもよい。なお、ここでいう「コンピュータシステム」とは、OSや周辺機器等のハードウェアを含むものとする。また、「コンピュータシステム」は、ホームページ提供環境(あるいは表示環境)を備えたWWWシステムも含むものとする。また、「コンピュータ読み取り可能な記録媒体」とは、フレキシブルディスク、光磁気ディスク、ROM、CD-ROM等の可搬媒体、コンピュータシステムに内蔵されるハードディスク等の記憶装置のことをいう。さらに「コンピュータ読み取り可能な記録媒体」とは、インターネット等のネットワークや電話回線等の通信回線を介してプログラムが送信された場合のサーバやクライアントとなるコンピュータシステム内部の揮発性メモリ(RAM)のように、一定時間プログラムを保持しているものも含むものとする。

【0048】

また、上記プログラムは、このプログラムを記憶装置等に格納したコンピュータシステ

ムかつ、広域ネットワークとして、あるいは、広域ネットワークの広域域により他のコンピュータシステムに伝送されてもよい。ここで、プログラムを伝送する「伝送媒体」は、インターネット等のネットワーク（通信網）や電話回線等の通信回線（通信線）のように情報を伝送する機能を有する媒体のことをいう。また、上記プログラムは、前述した機能の一部を実現するためのものであっても良い。さらに、前述した機能をコンピュータシステムにすでに記録されているプログラムとの組み合わせで実現できるもの、いわゆる差分ファイル（差分プログラム）であっても良い。

【図面の簡単な説明】

【0049】

【図1】本発明の一実施例による音響信号検出システムの構成例を示すブロック図である。

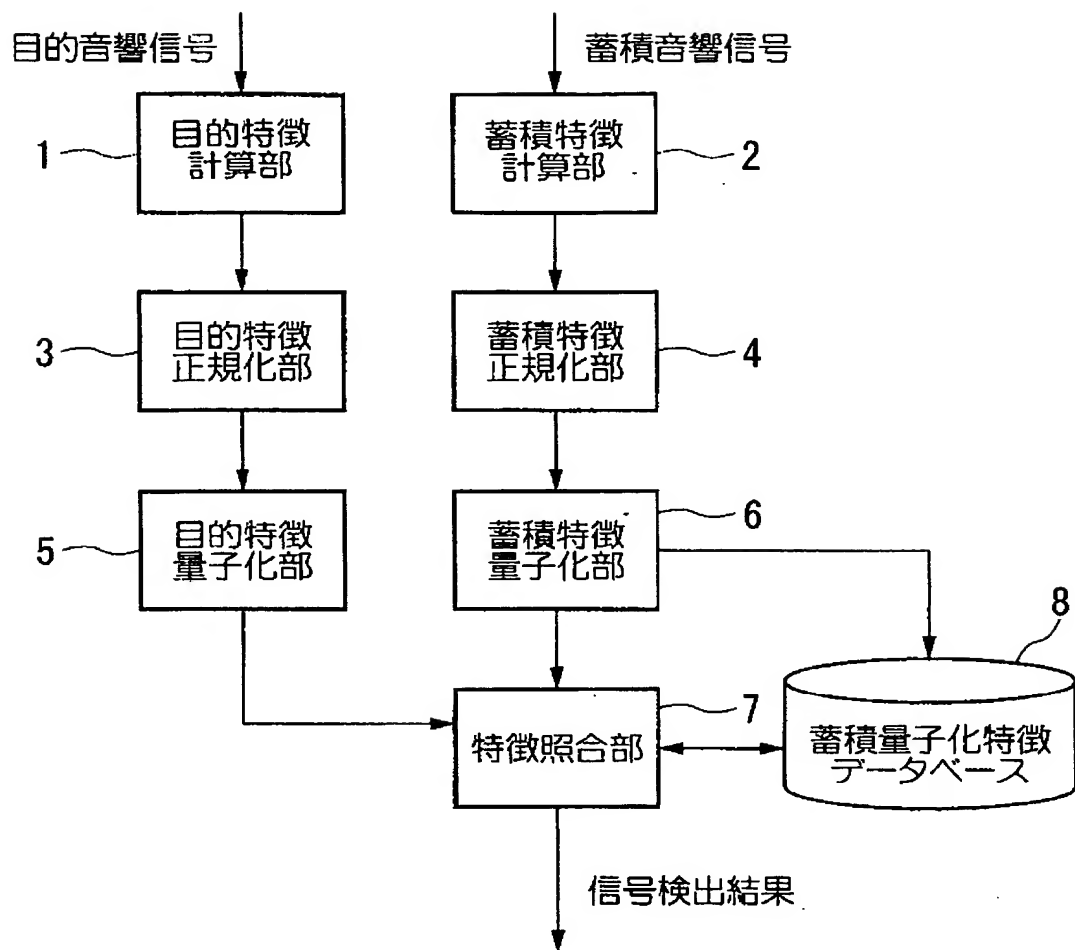
【図2】図1の音響信号検出システムの動作例を示すフローチャートである。

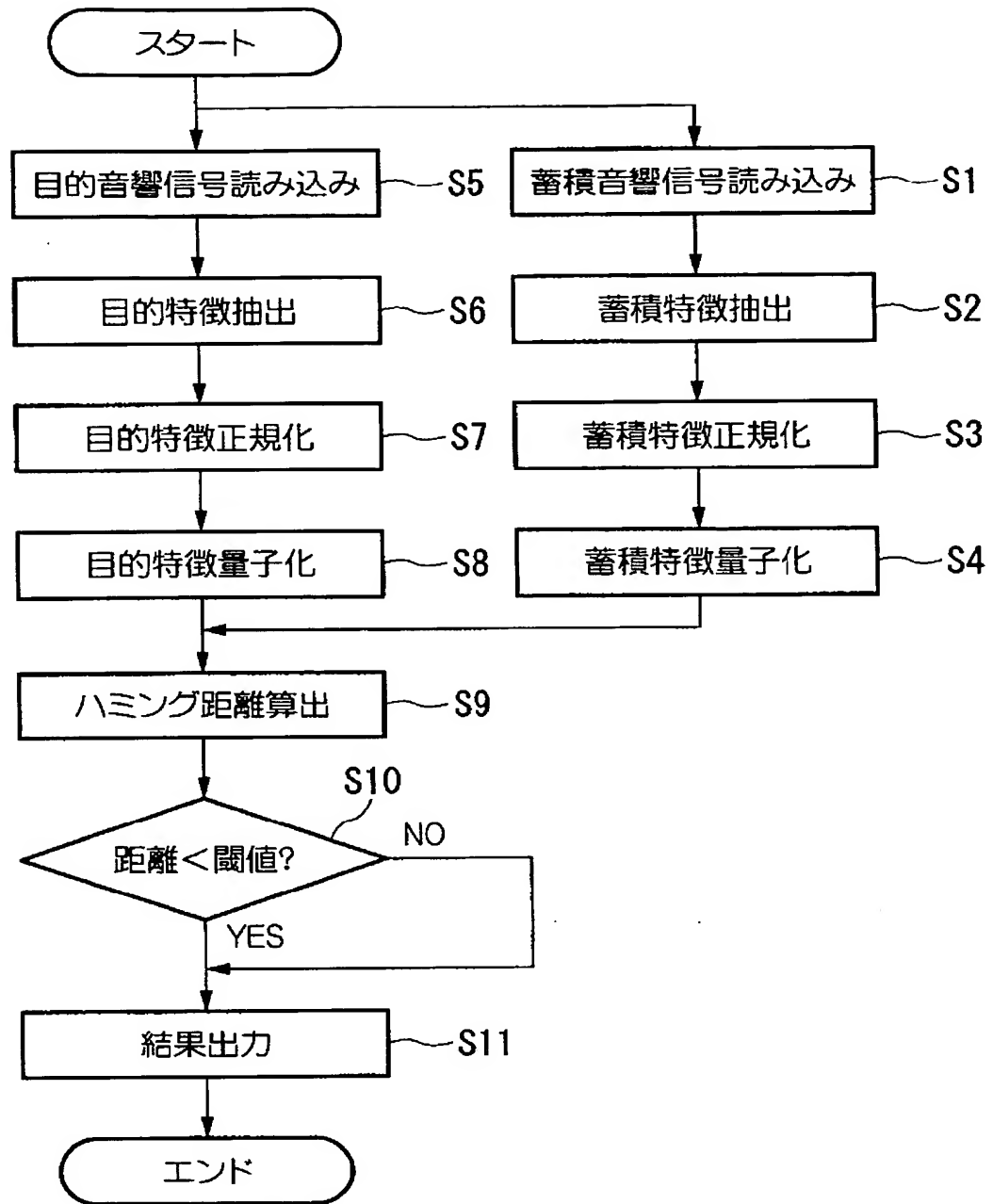
【符号の説明】

【0050】

- 1 … 目的特徴計算部
- 2 … 蓄積特徴計算部
- 3 … 目的特徴正規化部
- 4 … 蓄積特徴正規化部
- 5 … 目的特徴量子化部
- 6 … 蓄積特徴量子化部
- 7 … 特徴照合部
- 8 … 蓄積量子化特徴データベース

【 図 1 】





【要約】

【課題】 特徴ひずみに対して頑健なデータに変換して探索精度を向上させ、かつ探索時間を大幅に短縮させる音響信号検出システム及び音響信号検出サーバを提供する。

【解決手段】 本発明の音響信号検出システムは、蓄積音響信号から、目的音響信号に類似した部分を探すものであり、蓄積音響信号から蓄積特徴を計算する蓄積特徴計算部と、目的音響信号から目的特徴を計算する目的特徴計算部と、蓄積特徴を正規化して蓄積正規化特徴を導く蓄積特徴正規化部と、目的特徴を正規化して目的正規化特徴を導く目的特徴正規化部と、蓄積正規化特徴を量子化して蓄積量子化特徴を導く蓄積特徴量子化部と、目的正規化特徴を量子化して目的量子化特徴を導く目的特徴量子化部と、蓄積量子化特徴中に照合区間を設定し、前記目的量子化特徴及び前記蓄積量子化特徴中の該照合区間のそれぞれとの類似度を計算する特徴照合部とを有する。

【選択図】 図 1

0 0 0 0 0 4 2 2 6

19990715

住所変更

5 9 1 0 2 9 2 8 6

東京都千代田区大手町二丁目3番1号

日本電信電話株式会社

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record.

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☒ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.